

[2345/206]

## ENCRYPTION METHOD BASED ON FACTORIZATION

The present invention relates to an asymmetrical and public encryption method. In particular, the invention relates to a method for encrypting data on the basis of the factorization problem. In this context, the decryption of encrypted data is  
5 as complex as the problem of finding large prime divisors of large numbers. In detail, in the present invention, quadratic equations are to be solved for the decryption.

Encryption methods are used to protect data from unauthorized access when stored or during transmission over insecure  
10 communication channels. In so doing, the data are changed in such a way that this change cannot be undone without knowledge of a specific key. Encryption methods may be subdivided into the categories of asymmetrical and symmetrical. In symmetrical methods, the same key is used both for encryption and for  
15 decryption. Asymmetrical methods have two different keys, of which one is used for encryption and the other for decryption. In this context, all users can know the encryption key, whereas the decryption key must be kept secret. Therefore, the encryption key is also known as the public key, and the  
20 decryption key as the private key. Book [1] according to the literature list, for example, offers an overview of modern encryption methods.

The methods of Rabin ([3]) and Williams ([6]), which likewise utilize quadratic equations, are known. However, in these  
25 methods, only half the data bits is sent per transmission. Corresponding complexity restrictions thereby arise, and a greater demand for computing power during the encryption and the decryption.

Using polynomials of the second degree, the method of Schwenk and Eisfeld ([5]) offers little security against attacks which take advantage of the dependencies of message parts  $m_1$  and  $m_2$  on one another.

5 The objective is achieved by an invention having the features delineated in the independent claims. An asymmetrical encryption method is thereby described based on the factorization problem. It has less complexity than the RSA method in the encryption, and is able to transmit more data  
10 bits per encryption than the Rabin method or Williams method.

As already described above, the present invention concerns an asymmetrical encryption method. The public key is made up of a large composite number  $n$ ; the private key is made up of the factors of the composite number. The encryption is made up of  
15 a number of iterations of individual encryption steps that are successively reversed during the decryption. The reversal of an individual encryption step requires the solving of a quadratic equation modulo  $n$  (see below). Such a quadratic equation can only be easily solved if the factors of  $n$  are  
20 known.

The private key is preferably made up of the large prime numbers  $p$  and  $q$ . The public key is the product  $n$  of these two prime numbers, as well as a comparatively small integer  $L$  which is greater than one. Message  $m$  is made up of two  
25 integral values  $m_1$  and  $m_2$ , so that

$$m = (m_1, m_2),$$

both values lying in the set  $Z_n = \{0, 1, 2, \dots, n-1\}$ .

The encryption is accomplished via the equation

$$c = f^L(m) .$$

In the present case, encrypted value  $c$  is likewise made up of a double tuple of integers from  $Z_n$ , that is,  $c = (c_1, c_2)$ .

Function  $f^L(m)$  is recursively defined by

$$f^{j+1}(m) = f(f^j(m)).$$

5 For  $j = 1$ ,  $f^1(m) = f(m) = (f_1(m), f_2(m))$  applies, where

$$f_1(m) = m_1 + m_2 \bmod n$$

$$f_2(m) = m_1 \cdot m_2 \bmod n.$$

The encrypted text is therefore obtained by the recursions

$$a_{i+1} = a_i + b_i \bmod n \quad (1)$$

10 
$$b_{i+1} = a_i \cdot b_i \bmod n. \quad (2)$$

with the starting values  $a_0 = m_1$ ,  $b_0 = m_2$  and the final values  $c_1 = a_L$ ,  $c_2 = b_L$ .

For the decryption, one must be able to reverse the recursion. This is accomplished by solving the above equations for  $a_i$  and  
15  $b_i$ . One immediately obtains the quadratic equation

$$z^2 - a_{i+1} \cdot z + b_{i+1} = 0 \bmod n, \quad (3)$$

which has  $a_i$  and  $b_i$  as solutions. The problem of the further solutions of equation (3) will be discussed later. If  $n$  is the product of very large prime numbers, then the solution of  
20 quadratic equations without knowledge of the prime factors is presumably a very difficult problem. With knowledge of the prime factors, however, this is possible without difficulty. The current methods for taking the root modulo  $n$  are described in detail in [2].

25 To ensure the security of the encryption system, the recursion must be performed at least twice, since otherwise, if it is

performed exactly one time, the message parts  $m_1$  and  $m_2$  enter in linear fashion into the term  $a_1 = m_1 + m_2$ .

Another important aspect is the selection of the correct roots for the decryption.

- 5 If the number  $n$  contains exactly two prime factors  $p$  and  $q$ , equation (3) has four solutions. With a few bits for each  $a_i$ ,  $i = 1, 2, \dots, L$ , the sender is able to eliminate multivaluedness for the legitimate receiver. To resolve the multivaluedness, for example, error detection characters or  
10 parity characters may in each case be derived from  $a_i$ .

In the most favorable case, 2 bits per iteration step are needed to completely resolve the multivaluedness in each step. The 4 solutions of equation (3) are given by

$$z_{i,2,3,4} = \frac{a_{i+1}}{2} + w_{i,2,3,4} \bmod n \quad (4)$$

- 15 where

$$w_{i,2,3,4} = \sqrt{a_{i+1}^2 / 4 - b_{i+1}} \bmod n$$

are the four square roots of the above expression modulo  $n$ . The four values are connected as follows:

$$w_{i_1} = -w_{i_2} \bmod n \text{ and } w_{i_3} = -w_{i_4} \bmod n$$

- 20 We select the parity (even, odd) of the four roots so that

$$w_{i,3} = \text{even} \quad \text{and} \quad w_{i,4} = \text{odd}$$

One particularly elegant solution making it possible to differentiate all four roots from one another is as follows for  $p \equiv q \equiv 3 \bmod 4$ :

In addition to parity, the so-called Jacobi symbol  $(w_i/n)$  is used as a further discriminant criterion (for theory and efficient calculation, see, for example, [2]). For non-trivial values of  $w_i$ , as are needed in the decryption, the Jacobi symbol supplies the value 1 or -1. The Jacobi symbol can be calculated with expenditure  $O(\log^2 n)$ .

The parity and the Jacobi symbol are sufficient for precisely selecting one of the four roots  $w_{i,2,3,4}$ . The parity and the Jacobi symbol are able to be coded using 2 bits. By appending these two bits in each of the  $L$  iteration steps, the legitimate receiver is given the ability to reverse the  $L$  iteration steps.

The root leading to solution  $a_i$  in equation (4) is designated by  $w_i$ , thus,  $a_i = a_{i+1}/2 + w_i \bmod n$ . The parity and the Jacobi symbol are each specified with respect to this root. With the establishment of the value of  $a_i$ , the value for  $b_i$  then follows immediately as  $b_i = a_{i+1} - a_i \bmod n$ . In summary, one thus obtains

$$a_i = a_{i+1}/2 + w_i \bmod n \quad (5)$$

$$b_i = a_{i+1}/2 - w_i \bmod n. \quad (6)$$

In the encryption, at each step, from the number pair  $(a_i, b_i)$ , the pair  $(a_{i+1}, b_{i+1})$  is calculated, as well as the parity and the Jacobi symbol of  $w_i = (a_i - a_{i+1}/2) \bmod n$ .

With knowledge of the factorization, these steps can each be reversed by solving

$$\sqrt{a_{i+1}^2/4 - b_{i+1}} \bmod n,$$

parity and Jacobi symbol of this root being represented.

Another important aspect is the parameter selection. At present, realistic orders of magnitude for each of the two prime numbers are from approximately 510 bits, i.e.,  $n$  has a length of approximately 1020 bits. For  $L$ , a magnitude  
5  $O(\log \log n)$  is recommended; for  $n$  of 1000 bits, a value of approximately 3-10.

The bit lengths to be selected in the future may be oriented to the parameters of the RSA method.

An advantage of the method presented here is that the quantity  
10 of useful data is twice as great as in comparable methods.

Using standard algorithms, an encryption complexity of  $O(L \log^2 n)$  is reached, if one calculates the expenditure for a multiplication using  $O(\log^2 n)$ . When using current algorithms, one must reckon with an expenditure of  $O(L \log^3 n)$  for the  
15 decryption complexity. If an order of magnitude of  $O(\log \log n)$  is selected for  $L$ , a time advantage (in addition to the greater useful-data rate) results for the encryption compared to the RSA method.

As in the case of the Rabin method and Williams method, care  
20 must be taken in the implementation that, in each case, only the correct roots of equation (3) exit the decoder during the decryption, since otherwise the number  $n$  can be factored.

In another refinement, as in the RSA method, module [sic]  $n$  may also contain more than two large prime factors. Naturally,  
25 the number of solutions for equation (3) also increases accordingly.

A further generalization is achieved by introducing additional constants in the recursion:

$$a_{i+1} = k_1 \cdot a_i + k_2 \cdot b_i \bmod n$$

$$b_{i+1} = k_3 \cdot a_i \cdot b_i \bmod n,$$

which are made known as part of the public key. The decoding is performed in correspondingly modified form.

In another specific embodiment, the magnitude of the tuple is altered. Instead of working with double tuples  $m = (m_1, m_2)$ , it is also possible to work with  $q$  tuples. In the following, the expansion based on triple tuples is illustrated. The message is now made up of the triple tuple

$$m = (m_1, m_2, m_3).$$

10 The formula for the  $L$ th iteration step is still

$$f^{j+1}(m) = f(f^j(m)),$$

the basic iteration  $f^1(m) = (f_1(m), f_2(m), f_3(m))$ , however, being formed as follows:

$$f_1(m) = m_1 + m_2 + m_3 \bmod n$$

$$15 \quad f_2(m) = m_1 \cdot m_2 + m_1 \cdot m_3 + m_2 \cdot m_3 \bmod n$$

$$f_3(m) = m_1 \cdot m_2 \cdot m_3 \bmod n.$$

The inverse calculation is accomplished by solving a third-degree equation. The roots may again be discriminated by information (parity symbol, Jacobi symbol, etc.) derived accordingly from the interim results. The expansion to degrees greater than or equal to four may be accomplished in analogous manner. In the iteration, essentially the elementary-symmetric Newtonian terms must be considered, to which additional constants, as already described above, may be added.

25 In the following, the method of the present invention is elucidated in light of an example. For reasons of clarity, the numbers in the following are selected to be very small. Let us say  $n = 8549 = p \cdot q$ , with the private prime numbers  $p = 83$

and  $q = 103$ . Let us assume the number of iterations  $L = 3$ , and the message to be encrypted is given by  $m = (m_1, m_2) = (123, 456)$ . Even parity is coded by a zero, uneven parity by a one. Parity bit  $b_p$  is used for this. If the Jacobi symbol is equal to one, a one is coded, if it is equal to minus one, a zero is coded. Jacobi bit  $b_J$  is used for this.

The following values are obtained

$$(a_0, b_0) = (123, 456)$$

$$(a_1, b_1) = (579, 4794)$$

$$(a_2, b_2) = (5373, 5850)$$

$$(a_3, b_3) = (2674, 5926)$$

To each of the three pairs  $(a_1, b_1)$ ,  $(a_2, b_2)$  and  $(a_3, b_3)$ ,  $L \cdot 2$  bits of parity bits and Jacobi bits, given in the example by the following binary vector  $(b_{p_3}, b_{j_3}, b_{p_2}, b_{j_2}, b_{p_1}, b_{j_1}) = (0, 0, 1, 1, 0, 1)$ , are also added.

Initially, the receiver determines the four roots

$w_{2,1,2,3,4} = 1629, 4036, 4513, 6920$ . Based on  $b_{p_3} = 0$ , the receiver recognizes that the correct root is even. Thus, only 4036 and 6920 remain. Of these  $(4036/8549) = -1$  and  $(6920/8549) = 1$ .

$b_{j_3} = 0$  implies that 4036 is the correct selection. An analogous procedure leads to the complete decryption.

In certain application cases, e.g. when the unencrypted message  $m$  contains redundancy, it is possible to dispense with the co-transmission of the bits for resolving the

multivaluedness. For example, this is the case for normal texts or when a so-called hash value was already placed in  $m$ . However, this is done at a decryption expenditure increased by a factor of  $4^L$ . Corresponding compromises are likewise



possible; for example, the specification of only the parity in each of the  $L$  steps reduces the number of bits to be co-transmitted to  $L$  bits, and increases the decryption expenditure by the factor  $2^L$ .

- 5 As in the asymmetrical methods known in the literature ([1], [3], [4], [5]), a so-called digital signature method may be attained essentially by the interchange of encryption operations and decryption operations in the proposed method as well.

List of the Cited Literature:

- [1] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone,  
"Handbook of Applied Cryptography", CRC Press, 1996.
- [2] E. Bach, J. Shallit, "Algorithmic Number Theory", Vol. 1,  
5 Efficient Algorithms, The MIT Press, Cambridge, Massachusetts,  
1996.
- [3] M. O. Rabin, "Digitalized Signatures and Public-Key  
Functions as as [sic] intractable as Factorization ",  
MIT/LCS/TR- 212, 1979.
- 10 [4] R. L. Rivest, A. Shamir, L. Adleman, "A Method for  
Obtaining Digital Signatures and Public Key Cryptosystems",  
Communications of the ACM, Vol. 21 No.2, pp. 120-126, Feb. .  
1978.
- [5] J. Schwenk, J. Eisfeld, "Public Key Encryption and  
15 Signature Schemes based and [sic] Polynomials over  $Z_n$ ",  
Eurocrypt 1996, LNCS 1070, Springer-Verlag Berlin Heidelberg  
1996.
- [6] H. Williams, "A Modification of the RSA Public-Key  
Equation Procedure ", IEEE Transactions on Information Theory,  
20 Vol. IT-26, No. 6, November 1980.